

CLAIMS

What is claimed is:

- 1 1. An integrated voice activation detector for
- 2 detecting whether voice is present, the integrated voice
- 3 activation detector comprising:
 - 4 a semiconductor integrated circuit including,
 - 5 at least one signal processing unit to perform
 - 6 voice detection; and
 - 7 a processor readable storage means to store
 - 8 signal processing instructions for execution by the
 - 9 at least one signal processing unit to:
 - 10 detect whether noise is present to
 - 11 determine whether a noise flag should be set;
 - 12 detect a predetermined number of zero
 - 13 crossings to determine whether a zero crossing
 - 14 flag should be set;
 - 15 detect whether a threshold amount of
 - 16 energy is present to determine whether an
 - 17 energy flag should be set;
 - 18 detect whether instantaneous energy is

19 present to determine whether a instantaneous
20 energy flag should be set; and
21 utilize a combination of the noise,
22 zero crossing, energy, and instantaneous energy
23 flags to determine whether voice is present.

1 2. The integrated voice activation detector of
2 claim 1, wherein the signal processing instructions
3 further for execution by the at least one signal
4 processing unit to, perform fast Fourier transformation
5 (FFT) processing to determine whether a FFT flag should
6 be set.

1 3. The integrated voice activation detector of
2 claim 1, wherein the signal processing instructions
3 further for execution by the at least one signal
4 processing unit to, perform an interim voice activity
5 decision, a interim voice activity decision flag being
6 set to indicate voice has been detected by determining if
7 the instantaneous energy flag is set or the energy flag
8 is set and the noise flag is not set and the zero
9 crossing flag is not set.

042390.P12534

1 4. The integrated voice activation detector of
2 claim 3, wherein the signal processing instructions
3 further for execution by the at least one signal
4 processing unit to, perform HangOver and Speech Kick in
5 processing after the interim voice activity decision has
6 been made to determine whether a voice activity flag
7 should be set or cleared.

1 5. The integrated voice activation detector of
2 claim 4, wherein the signal processing instructions
3 further for execution by the at least one signal
4 processing unit to, if the voice activity flag is set,
5 send a speech payload to be packetized and update the
6 voice activity detection flag for external interaction
7 with other functions of the semiconductor integrated
8 circuit.

1 6. The integrated voice activation detector of
2 claim 4, wherein the signal processing instructions
3 further for execution by the at least one signal
4 processing unit to, if the voice activity flag is not

TOP SECRET//NOFORN

5 set, disable an automatic level control and cause a
6 silence insertion description payload to be prepared.

1 7. The integrated voice activation detector of
2 claim 1, wherein detecting a predetermined number of zero
3 crossings to determine whether a zero crossing flag
4 should be set includes determining whether a root mean
5 square crossing value is greater than a threshold value.

1 8. The integrated voice activation detector of
2 claim 1, wherein detecting whether noise is present to
3 determine whether a noise flag should be set includes
4 determining whether energy in a current frame multiplied
5 by a threshold is greater than delayed frame energy.

1 9. The integrated voice activation detector of
2 claim 1, wherein detecting whether a threshold amount of
3 energy is present to determine whether an energy flag
4 should be set includes determining if a logarithm of an
5 autocorrelation of a frame is greater than an energy
6 threshold.

042390.P12534

1 10. The integrated voice activation detector of
2 claim 1, wherein detecting whether instantaneous energy
3 is present to determine whether an instantaneous energy
4 flag should be set includes determining whether a
5 difference between a current frames energy at an
6 autocorrelation of a tenth delayed sample and a prior
7 frames energy at an autocorrelation of a tenth delayed
8 sample is greater than a previous frames autocorrelation
9 multiplied by a threshold.

1 11. A method for voice activation detection to
2 detect whether voice is present, the method comprising:
3 detecting whether noise is present to
4 determine whether a noise flag should be set;
5 detecting a predetermined number of
6 zero crossings to determine whether a zero
7 crossing flag should be set;
8 detecting whether a threshold amount of
9 energy is present to determine whether an
10 energy flag should be set;
11 detecting whether instantaneous energy

12 is present to determine whether a instantaneous
13 energy flag should be set; and
14 utilizing a combination of the noise,
15 zero crossing, energy, and instantaneous energy
16 flags to determine whether voice is present.

1 12. The method of claim 11, further comprising,
2 performing fast Fourier transformation (FFT) processing
3 to determine whether a FFT flag should be set.

1 13. The method of claim 11, further comprising,
2 performing an interim voice activity decision, a interim
3 voice activity decision flag being set to indicate that
4 voice has been detected by determining if the
5 instantaneous energy flag is set or the energy flag is
6 set and the noise flag is not set and the zero crossing
7 flag is not set.

1 14. The method of claim 13, further comprising,
2 performing HangOver and Speech Kick in processing after
3 the interim voice activity decision has been made to
4 determine whether a voice activity flag should be set or

5 cleared.

1 15. The method of claim 14, further comprising, if
2 the voice activity flag is set, sending a speech payload
3 to be packetized and updating the voice activity
4 detection flag for external interaction with other
5 functions.

1 16. The method of claim 14, further comprising, if
2 the voice activity flag is not set, disabling an
3 automatic level control and causing a silence insertion
4 description payload to be prepared.

1 17. The method of claim 11, wherein detecting a
2 predetermined number of zero crossings to determine
3 whether a zero crossing flag should be set includes
4 determining whether a root mean square crossing value is
5 greater than a threshold value.

1 18. The method of claim 11, wherein detecting
2 whether noise is present to determine whether a noise
3 flag should be set includes determining whether energy in

4 a current frame multiplied by a threshold is greater than
5 delayed frame energy.

1 19. The method of claim 11, wherein detecting
2 whether a threshold amount of energy is present to
3 determine whether an energy flag should be set includes
4 determining if a logarithm of an autocorrelation of a
5 frame is greater than an energy threshold.

1 20. The method of claim 11, wherein detecting
2 whether instantaneous energy is present to determine
3 whether an instantaneous energy flag should be set
4 includes determining whether a difference between a
5 current frames energy at an autocorrelation of a tenth
6 delayed sample and a prior frames energy at an
7 autocorrelation of a tenth delayed sample is greater than
8 a previous frames autocorrelation multiplied by a
9 threshold.

1 21. An apparatus comprising:
2 at least one signal processing unit to perform
3 voice detection; and

4 a storage device to store signal processing
5 instructions for execution by the at least one
6 signal processing unit to:

7 determine whether a noise flag, a zero
8 crossing flag, an energy flag, and an
9 instantaneous energy flag should be set; and
10 utilize a combination of the noise,
11 zero crossing, energy, and instantaneous energy
12 flags to determine whether voice is present.

1 22. The apparatus of claim 21, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to:

4 detect whether noise is present to
5 determine whether the noise flag should be set;
6 detect a predetermined number of zero
7 crossings to determine whether the zero
8 crossing flag should be set;
9 detect whether a threshold amount of
10 energy is present to determine whether the
11 energy flag should be set; and
12 detect whether instantaneous energy is

13 present to determine whether the instantaneous
14 energy flag should be set.

1 23. The apparatus of claim 21, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to, perform fast Fourier
4 transformation (FFT) processing to determine whether a
5 FFT flag should be set.

1 24. The apparatus of claim 21, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to, perform an interim
4 voice activity decision, a interim voice activity
5 decision flag being set to indicate voice has been
6 detected by determining if the instantaneous energy flag
7 is set or the energy flag is set and the noise flag is
8 not set and the zero crossing flag is not set.

1 25. The apparatus of claim 24, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to, perform HangOver and
4 Speech Kick in processing after the interim voice

5 activity decision has been made to determine whether a
6 voice activity flag should be set or cleared.

1 26. The apparatus of claim 25, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to, if the voice
4 activity flag is set, send a speech payload to be
5 packetized and update the voice activity detection flag
6 for external interaction with other functions of the
7 semiconductor integrated circuit.

1 27. The apparatus of claim 25, wherein the signal
2 processing instructions further for execution by the at
3 least one signal processing unit to, if the voice
4 activity flag is not set, disable an automatic level
5 control and cause a silence insertion description payload
6 to be prepared.

1 28. The apparatus of claim 22, wherein detecting a
2 predetermined number of zero crossings to determine
3 whether a zero crossing flag should be set includes
4 determining whether a root mean square crossing value is

5 greater than a threshold value.

1 29. The apparatus of claim 22, wherein detecting
2 whether noise is present to determine whether a noise
3 flag should be set includes determining whether energy in
4 a current frame multiplied by a threshold is greater than
5 delayed frame energy.

1 30. The apparatus of claim 22, wherein detecting
2 whether a threshold amount of energy is present to
3 determine whether an energy flag should be set includes
4 determining if a logarithm of an autocorrelation of a
5 frame is greater than an energy threshold.

1 31. The apparatus of claim 22, wherein detecting
2 whether instantaneous energy is present to determine
3 whether an instantaneous energy flag should be set
4 includes determining whether a difference between a
5 current frames energy at an autocorrelation of a tenth
6 delayed sample and a prior frames energy at an
7 autocorrelation of a tenth delayed sample is greater than
8 a previous frames autocorrelation multiplied by a

9 threshold.

1 32. A method comprising:

2 determining whether a noise flag, a zero
3 crossing flag, an energy flag, and an
4 instantaneous energy flag should be set; and
5 utilizing a combination of the noise, zero
6 crossing, energy, and instantaneous energy
7 flags to determine whether voice is present.

1 33. The method of claim 32,further comprising:

2 detecting whether noise is present to
3 determine whether the noise flag should be set;
4 detecting a predetermined number of zero
5 crossings to determine whether the zero
6 crossing flag should be set;
7 detecting whether a threshold amount of
8 energy is present to determine whether the
9 energy flag should be set; and
10 detecting whether instantaneous energy is
11 present to determine whether the instantaneous
12 energy flag should be set.

1 34. The method of claim 33, further comprising,
2 performing fast Fourier transformation (FFT) processing
3 to determine whether a FFT flag should be set..

1 35. The method of claim 32, further comprising,
2 performing an interim voice activity decision, a interim
3 voice activity decision flag being set to indicate that
4 voice has been detected by determining if the
5 instantaneous energy flag is set or the energy flag is
6 set and the noise flag is not set and the zero crossing
7 flag is not set.

1 36. The method of claim 35, further comprising,
2 performing HangOver and Speech Kick in processing after
3 the interim voice activity decision has been made to
4 determine whether a voice activity flag should be set or
5 cleared.

1 37. The method of claim 36, further comprising, if
2 the voice activity flag is set, sending a speech payload
3 to be packetized and updating the voice activity
4 detection flag for external interaction with other

PROSECUTOR'S COPY

5 functions.

1 38. The method of claim 36, further comprising, if
2 the voice activity flag is not set, disabling an
3 automatic level control and causing a silence insertion
4 description payload to be prepared.

1 39. The method of claim 33, wherein detecting a
2 predetermined number of zero crossings to determine
3 whether a zero crossing flag should be set includes
4 determining whether a root mean square crossing value is
5 greater than a threshold value.

1 40. The method of claim 33, wherein detecting
2 whether noise is present to determine whether a noise
3 flag should be set includes determining whether energy in
4 a current frame multiplied by a threshold is greater than
5 delayed frame energy.

1 41. The method of claim 33, wherein detecting
2 whether a threshold amount of energy is present to
3 determine whether an energy flag should be set includes

4 determining if a logarithm of an autocorrelation of a
5 frame is greater than an energy threshold.

1 42. The method of claim 33, wherein detecting
2 whether instantaneous energy is present to determine
3 whether an instantaneous energy flag should be set
4 includes determining whether a difference between a
5 current frames energy at an autocorrelation of a tenth
6 delayed sample and a prior frames energy at an
7 autocorrelation of a tenth delayed sample is greater than
8 a previous frames autocorrelation multiplied by a
9 threshold.

1 43. A machine-readable medium having stored thereon
2 instructions, which when executed by a machine, causes
3 the machine to perform operations comprising:

4 determining whether a noise flag, a zero
5 crossing flag, an energy flag, and an
6 instantaneous energy flag should be set; and
7 utilizing a combination of the noise, zero
8 crossing, energy, and instantaneous energy
9 flags to determine whether voice is present.

1 44. The machine-readable medium of claim 43, further
2 comprising:

3 detecting whether noise is present to
4 determine whether the noise flag should be set;
5 detecting a predetermined number of zero
6 crossings to determine whether the zero
7 crossing flag should be set;
8 detecting whether a threshold amount of
9 energy is present to determine whether the
10 energy flag should be set; and
11 detecting whether instantaneous energy is
12 present to determine whether the instantaneous
13 energy flag should be set.

1 45. The machine-readable medium of claim 43, further
2 comprising, performing fast Fourier transformation (FFT)
3 processing to determine whether a FFT flag should be set.

1 46. The machine-readable medium of claim 43, further
2 comprising, performing an interim voice activity
3 decision, a interim voice activity decision flag being
4 set to indicate that voice has been detected by

5 determining if the instantaneous energy flag is set or
6 the energy flag is set and the noise flag is not set and
7 the zero crossing flag is not set.

1 47. The machine-readable medium of claim 46, further
2 comprising, performing HangOver and Speech Kick in
3 processing after the interim voice activity decision has
4 been made to determine whether a voice activity flag
5 should be set or cleared.

1 48. The machine-readable medium of claim 47, further
2 comprising, if the voice activity flag is set, sending a
3 speech payload to be packetized and updating the voice
4 activity detection flag for external interaction with
5 other functions.

1 49. The machine-readable medium of claim 47, further
2 comprising, if the voice activity flag is not set,
3 disabling an automatic level control and causing a
4 silence insertion description payload to be prepared.

1 50. The machine-readable medium of claim 44, wherein

2 detecting a predetermined number of zero crossings to
3 determine whether a zero crossing flag should be set
4 includes determining whether a root mean square crossing
5 value is greater than a threshold value.

1 51. The machine-readable medium of claim 44, wherein
2 detecting whether noise is present to determine whether a
3 noise flag should be set includes determining whether
4 energy in a current frame multiplied by a threshold is
5 greater than delayed frame energy.

1 52. The machine-readable medium of claim 44, wherein
2 detecting whether a threshold amount of energy is present
3 to determine whether an energy flag should be set
4 includes determining if a logarithm of an autocorrelation
5 of a frame is greater than an energy threshold.

1 53. The machine-readable medium of claim 44, wherein
2 detecting whether instantaneous energy is present to
3 determine whether an instantaneous energy flag should be
4 set includes determining whether a difference between a
5 current frames energy at an autocorrelation of a tenth

6 delayed sample and a prior frames energy at an
7 autocorrelation of a tenth delayed sample is greater than
8 a previous frames autocorrelation multiplied by a
9 threshold.

00000000000000000000000000000000